

Use of a histopathology data pool for epidemiological analysis

JC MACARTNEY, TP ROLLASON, AND BW CODLING

From the Department of Pathology, The Medical School, Birmingham University, UK

SUMMARY Data from the Birmingham Histopathology Data Pool of surgical pathology on the age, sex, and site distributions of various cutaneous and hepatic lesions have been analysed, and the results have been compared with data from the Regional Cancer Registry. Analysis of malignant neoplasms shows that there is no significant difference between the two sources. The results are discussed with reference to the potential use of histopathology data banks for quality assurance in histopathology and epidemiological analysis.

The purpose of this study is to describe ways in which a histopathology data pool can be used to provide simple epidemiological information on disease patterns. Analysis of data pool material has been advocated as a means of measuring quality assurance in histopathology.¹ If such an approach is feasible, it is important to compare the consistency of the stored data with available sources of information. In this paper an attempt has been made to assess the quality of the stored data and to identify areas where improvements are required which would widen the potential value of such data pools.

Since its inception in 1970, the Birmingham Histopathology Data Pool (BHDP) has acquired a large volume of coded histological data, and significant numbers of both common and rare lesions are on file. Details of the project, which uses the Systematised Nomenclature of Pathology (SNOP)² for coding histopathological material have been described previously.³ For this study, coded data on several types of cutaneous neoplasms and various hepatic inflammatory lesions have been selected for analysis. The resulting information has been compared with available information from the Birmingham Cancer Registry.⁴ The results indicate that histopathology data pools have considerable potential with the advantage of providing information on benign neoplasms and non-neoplastic conditions.

Material and methods

The details of computer facilities, data input, processing, and routine output together with the

contributing laboratories have been described.³ For this study, retrievals of specific lesions were selected in either SNOP anatomical (skin neoplasms) or morphological (hepatic lesions) order. The cutaneous neoplasms, with equivalent SNOP morphological codes in parentheses, were squamous carcinoma (M8073), in-situ squamous carcinoma including Bowen's disease (M8072 and M8082), and various naevi and melanomas (M872- to M879-). The hepatic lesions selected were sarcoidosis (M4454), granulomatous inflammation not otherwise specified but excluding sarcoid (M44--), chronic inflammation (M43--), inflammation not otherwise specified, acute and subacute (M4--- to M42--), alcoholic cirrhosis (M4851 E5510), cirrhosis excluding alcoholic cirrhosis (M4850-M4861), and biliary cirrhosis (M4852). Retrievals were based on the whole data pool content for the years 1970 to 1977 inclusive. Multiple listings of the same lesion, coded as occurring at the same anatomical site in the same patient, were eliminated before analysis. Although data retrievals for skin neoplasms split the anatomical site of the lesions into multiple topographical areas the data have been regrouped into major sites for ease of presentation.

In a small proportion of cases of skin neoplasms the topographical site was not assigned and coded as skin, not otherwise specified (TO100). These cases are excluded from data on site distributions.

The following information was derived from the retrievals: age and sex distributions, and site distributions. The results for skin malignancies have been compared with published data from the Birmingham Regional Cancer Registry (BRCR)⁴ for the years 1957 to 1972.

Received for publication 8 October 1979

Table 1 Sex and age distribution of cutaneous neoplasms

	<i>Squamous carcinoma</i>		<i>Carcinoma in-situ</i>		<i>Basal cell carcinoma</i>		<i>Malignant melanoma</i>	
	<i>BHDP</i>	<i>BRCR</i>	<i>BHDP</i>	<i>BRCR</i>	<i>BHDP</i>	<i>BRCR</i>	<i>BHDP</i>	<i>BRCR</i>
<i>Males</i>								
Mean age (yr)	68.2	67.3	64.0	—	61.7	62.2	55.6	53.0
Median age (yr)	68.1	70.7	65.2	—	63.7	64.8	57.0	51.0
Standard deviation	10.4	12.8	—	—	11.8	11.1	15.2	16.9
95% between ages	51.89	42.89	40.83	—	37.85	36.87	26.82	21.82
Proportion by sex	65	68	41	—	54	52	29	34
Total cases	221	3416	28	—	695	9405	57	489
<i>Females</i>								
Mean age (yr)	71.8	70.2	63.0	—	65.0	65.6	52.7	53.1
Median age (yr)	76.1	74.5	65.0	—	65.5	69.5	57.0	50.0
Standard deviation	10.6	13.7	—	—	13.9	12.5	17.3	17.9
95% between ages	53.86	38.93	39.90	—	35.89	38.91	21.84	18.91
Proportion by sex	35	32	59	—	46	48	71	66
Total cases	120	1593	64	—	582	8564	141	949

Results

AGE AND SEX DISTRIBUTION OF CUTANEOUS NEOPLASMS

Table 1 shows the mean and median ages at diagnosis within the BHDP for each neoplasm analysed together with the age range within which 95% of cases fell, and the sex distribution expressed as a percentage of the total number of cases recorded. Similar data from the BRCR are shown alongside. Cases of carcinoma-in-situ of the skin and benign naevi are not documented by the Registry and no comparison figures are shown.

Figures for males and females are recorded separately. The mean ages of male and female cases of squamous and basal cell carcinoma in the BHDP figures are significantly different (for squamous carcinoma $0.003 > P > 0.001$ and for basal cell carcinoma $P < 0.001$). The mean ages of male and female cases of malignant melanoma, carcinoma-in-situ, and the various benign naevi were not significantly different at the 5% level of probability but there was a marked disproportion in the sex ratio, and figures have been grouped according to sex for ease of presentation.

The most striking feature of the data from the two

sources is that they are very similar. The mean ages of either male or female cases of squamous and basal cell carcinoma and malignant melanoma in the two series are not significantly different at the 5% level of probability. Comparison of the median ages, age ranges, and proportions of cases by sex shows that the data are similar.

SITE DISTRIBUTION OF CUTANEOUS NEOPLASMS

In view of the similar age and sex structures of the two series it is not surprising that the site distributions of the various neoplasms are similar. Table 2 shows the topographic distribution of squamous carcinoma. In both sources the upper limb in males and the lower limb in females are important sites for squamous carcinoma. This sex difference is further illustrated by examining the sex ratios of cases at different sites on the body (Table 2). Similar trends are shown in data from the BHDP on in-situ carcinoma (Table 3).

Table 4 shows the topographical distribution and sex ratios for basal cell carcinoma. Both series show similar features.

The anatomical distribution of cases of malignant melanoma together with sex ratios are shown in

Table 2 Distribution of cutaneous squamous carcinoma by site* and sex ratios

<i>Site</i>	<i>BHDP</i>			<i>BRCR</i>		
	<i>Males</i>	<i>Females</i>	<i>M:F ratio</i>	<i>Males</i>	<i>Females</i>	<i>M:F ratio</i>
Head and neck	64	53	2.3:1	65	55	2.2:1
Upper limb	24	9	4.6:1	24	14	3.1:1
Trunk	3	7	0.9:1	5	13	0.7:1
Lower limb	9	31	0.5:1	6	18	0.6:1

*Expressed as a percentage.

Site	Junctional naevus		Compound naevus		Intradermal naevus		Blue naevus		Juvenile melanoma	
	BHDP	BRCR	BHDP	BRCR	BHDP	BRCR	BHDP	BRCR	BHDP	BRCR
1	—	—	26.6	—	38.1	—	30.4	—	11.5	—
	—	—	25.0	—	36.0	—	30.0	—	10.0	—
	—	—	14.5	—	17.5	—	14.2	—	10.9	—
	—	—	7.60	—	13.73	—	14.52	—	2.20	—
	—	—	26	—	50	—	41	—	39	—
	—	—	121	—	244	—	19	—	13	—
2	—	—	27.7	—	36.8	—	30.1	—	13.9	—
	—	—	23.0	—	35.0	—	27.0	—	14.0	—
	—	—	13.0	—	15.4	—	14.1	—	8.5	—
	—	—	7.58	—	13.70	—	11.52	—	3.22	—
	—	—	74	—	50	—	59	—	61	—
	—	—	242	—	242	—	27	—	20	—

Table 5. Again, apart from the trunk in males, the data are similar and also show the importance of the lower limb in females as a site of occurrence.

Topographic distributions of four types of benign naevi recorded in the BHDP are shown in Table 6. All types of benign naevi showed a marked female predominance apart from intradermal naevi where the sex ratio was almost unity. The mean age of male and female cases of intradermal naevi was significantly higher ($p < 0.001$) than for other benign naevi. Interesting features from this table are the importance of the trunk and lower limbs as recorded sites of occurrence of junctional naevi and the trunk as a site for compound naevi. Blue naevi were recorded most frequently on the upper limb in females but, curiously, the head and neck was the most frequent site in males. Data for juvenile melanoma are not shown as numbers were small (see Table 1) but showed a slight tendency for the trunk to be the most frequent topographical site assigned. Other diagnoses recorded in the data pool included 12 cases of naevus not otherwise specified (M8720), nine cases of cellular naevi (M8790), and one malignant blue naevus recorded on the upper limb.

NON-ASSIGNED CASES

The number of cases where the anatomical site of a neoplasm was not assigned was small. For

squamous carcinoma the figures were 4% and 5% in males and females respectively. No cases of carcinoma-in-situ were unassigned. For basal cell carcinoma the figures were 1.7% and 3.6% and for malignant melanoma 1.8% and 2.8% respectively in males and females. Higher percentages of cases of benign naevi were not assigned specific topographical sites (circa 8%). Apart from junctional naevi in males, where 8% (2 cases) were unrecorded, the figures would not have seriously influenced the site distributions.

HEPATIC LESIONS

Hepatic inflammatory lesions were chosen for analysis as it was expected that the poor discriminating value of the inflammatory codes available in the SNOP coding system would not allow easy retrieval into homogenous groups. In practice, this proved to be the case. There was a wide scatter of ages among patients, and mean ages (data not shown) of different groups were not significantly different. With regard to cases of cirrhosis there was a similar wide scatter, and the only significant male excess was in cases coded specifically as alcoholic cirrhosis. Females were markedly in excess among cases coded as biliary cirrhosis, and the mean age for this diagnosis in males was rather higher than in females. However, wherever cases had relatively specific codes assigned to them the numbers were small. In the vast majority of cases the coding did not permit confident identification of the original morphological lesion described by the pathologist.

Discussion

SKIN NEOPLASMS

The age and site distributions and sex ratios in the BHDP for the three malignant neoplasms investi-

Table 3 Distribution of cutaneous squamous carcinoma in situ* by site and sex ratios

Site	Males	Females	M:F ratio
Head, neck, and upper limb	82	45	1.3:1
Trunk	7	7	0.7:1
Lower limb	11	48	0.2:1

*Data from BHDP only.

Table 4 *Distribution of cutaneous basal cell carcinoma by site and sex ratios*

Site	BHDP			BRCR		
	Males	Females	M:F ratio	Males	Females	M:F ratio
Head and neck	85	85	1.2:1	95	94	1.1:1
Upper limb	4	3	1.7:1	1	1	1.3:1
Trunk	8	8	1.3:1	3	4	0.7:1
Lower limb	3	4	0.9:1	1	1	0.8:1

Table 5 *Distribution of cutaneous malignant melanoma by site and sex ratios*

Site	BHDP			BRCR		
	Males	Females	M:F ratio	Males	Females	M:F ratio
Head, neck, and upper limb	32	28	0.5:1	46	30	0.8:1
Trunk	41	17	1:1	21	15	0.7:1
Lower limbs	27	55	0.2:1	33	55	0.3:1

Table 6 *Distribution by site of benign pigmented naevi*

Site	Junctional		Compound		Intradermal		Blue naevus	
	Males	Females	Males	Females	Males	Females	Males	Females
Head and neck	18	26	32	31	73	62	61	32
Upper limb	5	11	6	12	6	7	22	56
Trunk	36	31	48	46	19	27	11	8
Lower limb	41	32	14	11	2	4	6	4

gated are very similar to data from the Birmingham Regional Cancer Registry. This degree of concurrence is surprising and gratifying in view of the lack of a defined geographical population base for the sample population in the BHDP (see Discussion section on Accuracy of data). Both sources show a marked female preponderance on the lower limbs for cases of squamous carcinoma and malignant melanoma, reflecting the possible influence of sunlight exposure on the development of these tumours. Figures from the BHDP also illustrate the importance of the lower limb as a site for in-situ carcinoma contrasted with the small percentage of cases of basal cell carcinoma recorded at this site. The validity of treating cases of in-situ carcinoma as a single lesion may be disputed but no obvious differences emerged between cases coded as 'in-situ carcinoma' (M8072) and Bowen's disease (M8082).

Overall our data for malignant melanoma are very similar to that described for series from South London⁵ and Glasgow,⁶ although the latter series differs with regard to the site distribution in males.

As far as we are aware, no comparable data are available for benign pigmented naevi. In one published series⁷ of excised naevi, 50% occurred on the head and neck and 16% on the lower limbs. However, no information was given on the sex distribution of the series nor were the lesions categorised into different histological subtypes.

Obviously the figures of surgically excised naevi from the BHDP in no way reflect the overall incidence of these lesions in the general population, and the removal of naevi for cosmetic reasons is likely to influence sampling. Despite this reservation it is interesting to note that there is a tendency for distinct topographic patterns of distribution to occur for each histological subtype. While the degree of histological sampling may have influenced the precise diagnosis made in individual cases, it is likely that the figures do reflect the dominant patterns present in terms of the relative importance of junctional or intradermal components. Supporting this interpretation is the fact that the mean age at diagnosis of cases of intradermal naevi is significantly higher than that of other naevi. There was no evidence of bimodality in the age distributions of the benign naevi. We have no explanation for this other than the postulated influence of hormonal factors on junctional activity.

HEPATIC LESIONS

These data were chosen for analysis as they provide a different level of diagnostic activity and accuracy from that of the cutaneous neoplasms. The data proved difficult to interpret for two reasons. Firstly, the diagnostic criteria for hepatic inflammatory lesions are less easy to define and apply than those for many neoplasms. Secondly, the morphological

codes for inflammation provided in the SNOP code do not allow accurate retrospective analysis of cases into recognised disease categories. The plethora of inflammatory codes available in SNOP probably discourages the pathologist from discriminating selection. Specific codes for different disease entities or agreed coding conventions between different laboratories must be used if storage of this type of data is to be of any use.

ACCURACY OF DATA

By comparison with Cancer Registry figures the data on skin neoplasms appear to give a relatively consistent picture and by that criterion they appear accurate. We have previously estimated that an overall maximum potential coding error in the region of 8% could occur. This error comprised: (1) typographical errors (1%) due to misreading poor handwriting and the transposition of digits, (2) miscoding by the pathologist (3%), and (3) coding similar lesions in different ways due to inherent choice in the SNOP system (3%). By choosing well-defined cutaneous morphological entities for our analysis, errors due to the third factor are likely to be low. In contrast, the error rate in the hepatic lesions due to coding choice is likely to be much higher for the reasons discussed in the previous section.

In so far as all histological diagnoses made on surgical biopsy specimens examined in contributing laboratories are stored in the data pool, the information is complete. However, in two areas the data are deficient.

The first deficiency relates to the proportion of cases of cutaneous neoplasms where a topographical site was not assigned. For most of the neoplasms these cases constituted a small proportion of the overall numbers and probably would not significantly affect data on site distribution. The exceptions were cases coded as junctional naevi where 8.4% of male cases were not assigned specific topographic codes. This may have distorted the site frequency.

The second deficiency in our data is more major and relates to the potential use of histopathology data pools for providing incidence figures of lesions in the population. In contrast to the Cancer Registry, the Data Pool as constituted at present does not collect information from a defined geographical region. Despite the lack of a defined population base comparison between the sources shows a large measure of agreement. Apart from quality assurance,

histopathology data pools could give potentially valuable epidemiological information on diseases and be used to assess the impact of new investigative techniques on rates of diagnosis. With the decline in necropsy rates, this type of approach could be of immense value in assessing trends in both neoplastic and non-neoplastic conditions, provided current limitations can be overcome. If a histopathology data bank is to fulfil this potential, it is apparent that it should serve a defined hospital or geographical population where possible, and it is important that coding practice should be made uniform.

We are indebted to Professor Curran and the consultant histopathologists who contribute to the Birmingham Histopathology Data Pool who have made this study possible. Our thanks are also due to Mr BF Scott and Mr S Sargent, of the Queen Elizabeth Medical Centre Computer Unit, to Dr JAH Waterhouse for permission to use figures from the Birmingham Cancer Registry, to Mr A Minawu, of the Cancer Registry, for help with statistics, and to Miss G Parkinson for typing the manuscript.

References

- ¹ Henson DE, Codling BW, Macartney JC. Inter-laboratory Histological Evaluation: A new approach to quality control in Anatomic Pathology. Skokie, Illinois: College of American Pathologists, 1976.
- ² College of American Pathologists, Committee on Nomenclature and Classification of Disease. *Systematized Nomenclature of Pathology*, Chicago, 1965.
- ³ Codling BW, Alexander MK, Parker RGF, Curran RC. The Birmingham Histopathology Data Pool: a co-operative project among 10 laboratories. *J Clin Pathol* 1977;**30**:1110-8.
- ⁴ Waterhouse JAH. *Cancer Handbook of Epidemiology and Prognosis*. Edinburgh and London: Churchill Livingstone, 1974.
- ⁵ Raven RW. The clinico-pathological aspects of malignant melanoma. *Ann NY Acad Sci* 1963;**100**: 142-65.
- ⁶ Cochran AJ. Malignant melanoma. A review of 10 years' experience in Glasgow, Scotland. *Cancer* 1969; **23**:1190-9.
- ⁷ Pack GT, Lenson N, Gerber DM. Regional distribution of moles and melanomas. *Arch Surg* 1952;**65**:862-70.

Requests for reprints to: Dr JC Macartney, Department of Pathology, The Medical School, Birmingham B15 2TJ, UK.